ORIGINAL RESEARCH PAPER

Computer Science

COLLABORATIVE FILTERING BASED WEB SERVICE RECOMMENDER SYSTEM USING USERS' SATISFACTION ON QOS ATTRIBUTES

KEY WORDS: Web service, Collaborative filtering, Non-Functional attributes, Pearson CorrelationCoefficient

G. Vadivelou

Department of Computer Science, Kanchi Mamunivar Centre for PG Studies, Puducherry.

ABSTRACT

To select the best service meeting their requirements, service users prefer to personalize their non-functional attributes, such as reliability and price. But, that creates challengesas service providers have to considercontradictory non-functional attributes in the service selection process for service users. Present memory-based collaborative filtering(CF) service recommendation methods employs this recommendation technique using non-functional attribute values obtained at service invocation in order to calculate the similarity between users or items, and also to predict missing non-functional attributes. However, this approach is not sufficient as the non-functional attribute values of invoked services may not necessarily satisfy their individual preferences. In this paper, a CF-based service recommendation method using users' individual preference on non-functional attributes is proposed.

I.INTRODUCTION

Services technology is well recognized as an easy way to integrate applications without boundaries. Hence, most organizations publish their services on the web for easy consumption by the public. This has increased the number of available services. Different service providers may offer similar services which provide same functionality. Hence, users are facing challenges inselecting the services which satisfies their requirements. But, users are showing much interest to non-functional attributes such as reliability, availability, response time etc., rather than selecting services based only on their functionality. The reason for this is the non-functional attributes give a difference among the similar services; enabling users to select the services which best outfit theirrequirements.

Collaborative filtering (CF) is one of the widely used service recommendation techniques that bases its recommendations on the ratings or behavior of other users in thesystem [1] and it assumes that, if users agree about the quality or relevance of some service items, then they will likely agree about other service items as well. Memory-based CF methods achieve this by calculating the similarity values between users or service items using non-functional attribute values collected when the service is invoked. The non-functional attribute values observed by users when the service is invoked may not resemble their satisfaction regarding that service. Hence, without taking into account the personalized preferences of users while computing similarity sometime may create a space between users' non-functional attribute value and their liking.

Hence in this paper, a new approach considering users' individual preferences, along with the non-functional attribute values of invoked services, to precisely recommend services to anactive/service user is proposed. The proposed method, accurately compute the similarity between users or service items by incorporating users' individual preferences on non-functional attributes in the proposed similarity function. The improved similarity function at first discovers whether the two service users or items share some past experiences. If yes, the Pearson Correlation Coefficient (PCC) [1, 2, 3& 4] is modified to consider liking of users' individual preferences on non-functional attributes. If not, the similarity between the user's preferences is computed. Based on the similarity values, the top-k algorithm is used to locate similar neighbors. Finally, to predict missing non-functional attribute values, the weighted average with mean offset is extended to include users' satisfaction on non-functional attributes based on their individual preferences.

II. BACKGROUND AND RELATED WORKS

There are limited research works that uses memory-based CF methodsto service recommendation. While some of these works used either user-based CF or item-based CF, others focus on hybrid memory based CF methods (a combination of theuser-based and item-based CF). Zheng et al. [4] proposes a hybrid CF algorithm for recommending web services. Sreenath and Singh [5] and Rong et al. [6] use the idea of CF in their work, and used MovieLens data [54] for experimental analysis. Shao et al. [7] uses a user-based CF algorithm to predict QoS values. These research works neither considered users' individual preferences on QoS and hencethe accuracy of the predicted values of these methods was disappointing.Personalized servicerecommendation has beenstudied in recommendation systems. Chen et al. [3] uses a personalized QoS-aware recommendation method considering the QoS variance based on users' locations in order to recommend services. Their approach assumes that users who are closely located with each other are more likely to have similar service experience compared to those who live far away from each other.Shao et al. [7] proposes an approach for personalized QoS prediction for web services via CF that considers the different experiences of users on the quality of the same web service. Their approach predicts QoS for web services, taking the similarity among consumers' experiences into consideration. Their assumption was that consumers, who have similar historical experiences on some services, would have similar experiences on other services. Jiang et al. [2] proposes ahybrid personalized CF-based recommendation method considering the contribution ofan object (service item) to the similarity degree between users. Their method considers that if two users invoked the same service item in the past, it does notpromise that those users are similar. In their work, they calculated the contribution of aservice item by computing the standard deviation of the QoS metrics for the service item.

Though the above approaches aims at personalizing service recommendation either through location or user experiences, they does not consider users individual preferences on non-functional attributes for recommending services. Because of this issue, these approaches for recommending services suffer from low prediction accuracy. Hence, this paper proposes an innovative CF algorithm for recommending services considering users' individual preferences on non-functional attributes and experiments conducted with real-world data proved that the proposed method performs better than the existing method.

III. PROPOSED RECOMMENDATION ALGORITHM CONSIDERING INDIVIDUAL PREFERENCES

Prior to recommending services, it is necessary to know the history of an active user with respect to his/her non-functional attribute. This is important because non-functional attribute information plays part in making accurate service recommendations. Due to this, our method collects historical non-functional attribute record of active users and stores this information in the non-functional attribute values history repository. Besides the non-functional attribute values history, it is also necessary to obtain the active user's personalized preference in order to personalize the services recommended to him/her. The personalized preference component is used tocollect this information.

Using both the non-functional attribute historical data and the personalized preference of the active user, the satisfaction of this user can be computed for each service in the service repository. Based on the satisfaction of the services, the similarity between users can then be computed and subsequently similar users (in case of user-based personalized preference recommendation) or similar items (in case of the item-based personalized preference recommendation) can be identified. Once the similar users and/or similar items are obtained, the respective missing values of the active user are predicted. Finally, the recommender weighs the two predicted values to recommend optimal services to the active user.

The proposed individual preference recommendation algorithm is formulated by using the memory-based CF method. This section discusses the different aspects of the algorithm.

Most of the recommender systems use Pearson Correlation Coefficient (PCC) [1, 2, 4] approach for computing similarities between users/items. But this approach suffers from some limitations as it uses only the non-functional attribute values for finding similarities considering the users' personalized preferences, (i.e.) PCC approach over estimates the similarities among users and/or items, especially, when co-invoked services or common users [4] are very less. Zheng et al. [4] addressed this issue by introducing a similarity weight for reducing the influence of very less number in similar co-invoked items. But, that approach does not consider individual preferences of users which create a gap between users' non-functional attribute values and theirsatisfaction on the services.

Also, PCC strongly depends on the commonly invoked services between users for calculating similarities and it assumes that they are not similar when the two users or service items have no co-invoked services. As this might be true for some users, it is not always correct, especially for users who has no service invocation history.

Hence, in order to overcome the above mentioned limitations of PCC approach for similarity computation between anytwo service users or service items, this work proposes an approach for similarity computation considering the users individual preference on the non-functional attribute which is given by the user as part of the query.

The degree of similarity between these two users based on the irindividual preference *Sim Per* (a,b) is defined as:

$$Sim_{Per}(a,b) = \delta * Sim_{Pre}(a,b) + (1-\delta) * Sim_{Sat}(a,b) \quad (1)$$

where Sim_{Pre} (a,b)is the degree of similarity between the individual preferences ofusers a and bif there are no co-invoked services between them, Sim Sat (a,b)is the degreeof similarity of users a and b based on the satisfaction of their individual preferences,if they have co-invoked service items, and δ is a adjustable parameter which decideswhich method to use.

Given two service users *a*and b, their individual preferences P_a^i and P_b^ion some non-functional attribute *i* and a distance measure *D*, the degree of similarity between the overall preference of users *a* and *b*,Sim_{PPre} (a,b), using distance based assessment proposed by Koczy [8] is given by:

$$Sim_{PPre}(a,b) = \frac{1}{m}\sum_{i=1}^{m}\frac{1}{1+D(P_a^i,P_b^i)} \quad (2)$$

where the normalized Hamming distance[8] is considered for D.

Similarity between satisfactions of user'sindividual preferences
Let *a* and b be two service users and $P_a$and P_b be the individual preference of both users respectively. The degree of similarity between these two users, based on the satisfaction of their individual preferences, Sim_{PS} (a,b) is computed as:

$$Sim_{PS}(a,b) = \frac{\sum_{i\in I}\left(Sat_{P_a}(S_i)-\overline{Sat_{P_a}}\right)\left(Sat_{P_u}(S_i)-\overline{Sat_{P_b}}\right)}{\sqrt{\sum_{i\in I}\left(Sat_{P_a}(S_i)-\overline{Sat_{P_a}}\right)^2}\sqrt{\sum_{i\in I}\left(Sat_{P_b}(S_i)-\overline{Sat_{P_b}}\right)^2}}$$

(3)

where, $I= I_a \cap I_b$ is the set of co-invoked service items by both users *a* and*b*, $Sat_{Pa}(S_i)$ and $Sat_{Pu}(S_i)$are the respective satisfaction degrees of service itemS_i based on the individual preference of users *a* and *b*, and Sat_{Pa} and Sat_{Pb} represent the mean satisfaction degrees based on the individual preference of users *a* and *b* respectively.

Predicting values for missing satisfaction: With similar neighbors, *N*, of theactive user identified, predictions for an active user's non-functional attribute value canbe generated for a service item *iS*. This is done by combining the satisfaction values ofusers in *N*. This is typically done by computing the weighted average with mean offset [1, 2, 4] of the neighboring users. This function is extended to compute the weighted meanoffset of the satisfaction values of users in *N* using the computed similarity values asweights.

The predicted satisfaction value for a service item S_i, $Sat_{Pa}(S_i)$, for an active user ausing the degree of similarity between users, based on the satisfaction of theirindividual preferencesis given as follows:

$$Sat_{P_a}(S_i) = \overline{Sat_{P_a}} + \frac{\sum_{u\in N} Sim_{Pre}(a,b)(Sat_{P_b}(S_i)-\overline{Sat_{P_b}})}{\sum_{b\in N} Sim_{Pre}(a,b)} \quad (4)$$

where Sat_{Pa} is the vector of average satisfaction value of different services based on the individual preference of the active user P_a, and (Sat_b) is the vector of averagesatisfaction value of different services based on the individual preference of the similar service user P_b.

Similarly, the satisfaction value for a service item S_i, $Sat_{Pa}(S_i)$, of an active user *a*, is predicted using the degree of similarity between service items, based on the satisfaction of their individual preferences as follows:

$$Sat_{P_a}(S_i) = \overline{Sat_{P_a}} + \frac{\sum_{u\in N} Sim_{Pre}(i,j)(Sat_{P_u}(S_i)-\overline{Sat_{P_b}})}{\sum_{u\in N} Sim_{Pre}(i,j)} \quad (5)$$

where Sat_{Pa} is the vector of average satisfaction value of different services based on the personalized preference of the active user P_a, and Sat_{Pb} is the vector of average satisfaction value of different services based on the individual preference of the similar service user P_b.

Finally, in order for recommending service(s) to the active user, the two predicted satisfaction values (satisfaction values from user-based and item-based) are combined. Since these two predicted values may have different prediction performance, the tunable parameter method [2, 4] was adopted tocombine the two values using the equation below:

$$Sat_{final} = \mu * Sat_{user} + (1-\mu) * Sat_{item} \quad (6)$$

where μ is the tunable parameter which determines which method to use (either the user-based,the item-based or both).Based on the predicted satisfaction values, services are recommended to the active user.

IV. IMPLEMENTATION AND OBSERVED RESULT

The QWS real-time dataset released by the author [7] is considered for experiment which consists of 2,507 records. For each record in the dataset,11 QoS parameters exist. As the selected dataset doesn't include user preferences, they were generated at random.

The focus is to have very sparse dataset matrices in order to see how the proposed recommendation system works on users with no or few co-invoked services.For this, 95%, 90% and 85% of response time and throughput values were removed at random from the dataset and sparse matrices were generated with density 5%, 10% and 15%, respectively.

For each non-functional attribute, the number of values made available toactive users were varied from 10 and 20, and name them Given 10, and Given 20, respectively. Then the proposed method for predicting themissing satisfaction values is used and finally services are recommended to the service user.

To validate the prediction performance of the proposed approach, the well-known hybrid recommendation methods, WSRec[4] is also implemented and the results were compared. The Normalized

Mean Absolute Error (NMAE) was used to measure the prediction accuracy.

Table 1.1 and 1.2 shows the comparison of proposed to WSRec approach on the response-time and throughput non-functional attribute respectively. The tables prove that the proposed method produces a smaller NMAE compared to the WSRec method for all 5%,10% and 15% densities. Hence, the prediction accuracy of the proposed approach is better than our baseline approach (WSRec).

To study the impact of the parameter μ to the proposed collaborative filtering method, the Top-K value was set to 10 and varies the value of μ from 0 to 1 with a step value of 0.1. Figure 1.1 and 1.2 shows the results of given number = 10 and 20 with 15% data matrix density for response time and throughput attributes respectively.

V. CONCLUSIONS AND FUTURE WORKS

This paper proposed an innovative method for service recommendation considering the individual preference of users. To accurately compute the similarity between users or service items, the proposed method extends the PCC method to incorporate satisfaction of users' individual preferences on non-functional attributes. Based on the similarity values, the Top-K method was employed to find similar neighbors. Finally, for predicting missing non-functional attribute values, the weighted average with mean offset is extended to incorporate users' satisfaction on non-functional attributes based on their preferences. It is observed that the proposed approach show

TABLE 1.1: Comparison of WSRec and proposed approach on Response Time attribute

Density	5%		10%		15%	
Given number	10	20	10	20	10	20
Baseline Approach [WSRec]	0.6387	0.6092	0.5789	0.5498	0.5248	0.4978
Proposed Approach	**0.3142**	**0.3012**	**0.2868**	**0.2984**	**0.2674**	**0.2622**

TABLE 1.2: Comparison of WSRec and proposed approach on Throughput attribute

Density	5%		10%		15%	
Given number	10	20	10	20	10	20
Baseline Approach [WSRec]	0.7868	0.7598	0.9028	0.8762	0.9341	0.9089
Proposed Approach	**0.3878**	**0.3921**	**0.4212**	**0.4172**	**0.4101**	**0.4117**

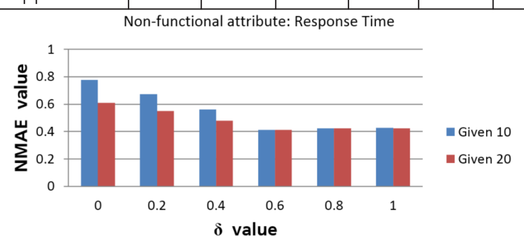

Fig. 1.1: NMAE values varying δ value (Response Time attribute)

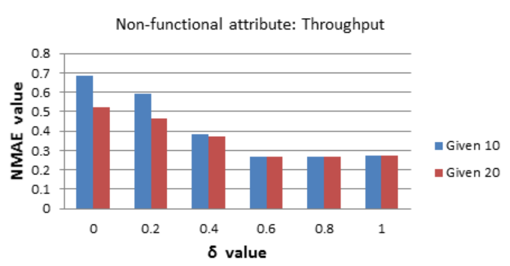

Fig. 1.2: NMAE values varying δ value (Throughput attribute)

significant improvement in the prediction accuracy compared to the existing method.

In web service selection process, different users may follow different decision making strategies, some are compensatory and some are non-compensatory. The proposed work doesn't consider the decision strategies in the ranking process, which are crucial or generating accurate ranking results for individual users. Hence, in the future,decision strategy may be incorporated in the ranking process. Also, the decision strategy could be integrated into some state-of-the-art selection models such as CP, AHP as the base selection algorithms instead of just considering some simple ranking rules.

REFERENCES

[1]. M.D. Ekstrand, J. T. Riedl, and J. A. Konstan. 2011. Collaborative filteringr ecommendersystems. Foundations and Trends in Human-ComputerInteraction 4, 2, 81-173.

[2] Y. Jiang, J. Liu, M. Tang, and X. Liu. 2011. An Effective Web Service Recommendation Method Based on Personalized CollaborativeFiltering. InProceedings ofthe 9th International Conference on WebServices (ICWS'11)..IEEE,Los Alamitos, CA, 211-218.DOI: 10.1109/ICWS.2011.38.

[3] X. Chen, Z. Zheng, X. Liu, Z. Huang, and H. Sun. 2013. PersonalizedQoSaware Webservice recommendation and visualization. IEEETransactions on Service Computing. 6, 1, 35–47. DOI:10.1109/TSC.2011.35.

[4] Z. Zheng, H. Ma, M.R. Lyu, and I. King. 2009. WSRec: A CollaborativeFiltering Based Web Service Recommender System. InProceeding ofthe 7th International Conference onWeb Services (ICWS'09). IEEE, LosAlamitos, CA, 437–444. DOI:10.1109/ICWS.2009.30.

[5] R.M. Sreenath and M.P. Singh. 2003. Agent-Based ServiceSelection.Journal WebSemantics. 1, 3, 261–279.DOI:10.1016/j.websem.2003.11.006.

[6] W.Rong, K. Liu, and L. Liang. 2009. Personalized Web ServiceRanking via User Group Combining Association Rule. In Proceedingsofthe 7thInternational Conference on Web Services (ICWS'09). IEEE,Los Alamitos, CA, 445–452. DOI: 10.1109/ICWS.2009.113.

[7] L. Shao, J. Zhang, Y. Wei, J. Zhao, B. Xie and H. Mei. 2007.Personalized QoS Prediction for Web Services via Collaborative Filtering. InProceedings of the 5th International Conferenceon Web Services (ICWS'05). IEEE, Los Alamitos, CA, 439–446. DOI:10.1109/ICWS.2007.140.

[8] Ismat Beg and Samina Ashraf. 2009. Similarity Measures for Fuzzy Sets.